STA 104 Solutions

Homework 6

Due Oct 13, 2005

§ 3.2

Problem 19, *p.* 184

Denote R, B, W and G as the events of choosing a red, blue, white and green ticket respectively, then

$$r = \mathsf{P}[R] = \frac{2}{7}, \quad b = \mathsf{P}[B] = \frac{1}{7}, \quad w = \mathsf{P}[W] = \frac{1}{7}, \quad g = \mathsf{P}[G] = \frac{3}{7}$$

a) Since there are at most 4 different colors,

$$\begin{aligned}
\mathsf{P}[X \geq 4] = \mathsf{P}[X = 4] &= \frac{5!}{2!1!1!1!} \times (r^2bwg + rb^2wg + rbw^2g + rbwg^2) \\
&= 3! \times \binom{5}{2} \times rbwg(r + b + w + g) \\
&= 0.1499
\end{aligned}$$

b) Suppose I_R, I_B, I_W and I_G are indicate functions, then

$$\begin{aligned}
\mathsf{E}[X] &= \mathsf{E}[I_R + I_B + I_W + I_G] = \mathsf{E}[I_R] + \mathsf{E}[I_B] + \mathsf{E}[I_W] + \mathsf{E}[I_G] \\
&= \mathsf{P}[I_R] + \mathsf{P}[I_B] + \mathsf{P}[I_W] + \mathsf{P}[I_G] \\
&= (1 - (1 - r)^5) + (1 - (1 - b)^5) + (1 - (1 - w)^5) + (1 - (1 - g)^5) \\
&= (1 - (\frac{5}{7})^5) + 2 \times (1 - (\frac{6}{7})^5) + (1 - (\frac{4}{7})^5) \\
&= 2.8278
\end{aligned}$$

Problem 20, *p.* 184

Suppose $\mathsf{P}[X = k] = p_k$, $k = 0, 1$ and 2. Then we have following equations,

$$\begin{cases}
\mathsf{E}[1] = & 1 \times p_0 + 1 \times p_1 + 1 \times p_2 = \quad 1 \\
\mathsf{E}[X] = & 0 \times p_0 + 1 \times p_1 + 2 \times p_2 = \quad \mu_1 \\
\mathsf{E}[X] = & 0 \times p_0 + 1 \times p_1 + 4 \times p_2 = \quad \mu_2
\end{cases}$$

Since the determinant of the coefficients in this system of linear equation is not zero, the distribution of X is determined. Solve it, we have

$$\begin{cases} p_0 = & 1- & \frac{3}{2}\mu_1 + & \frac{1}{2}\mu_2 \\ p_1 = & & 2\mu_1 - & \mu_2 \\ p_2 = & & -\frac{1}{2}\mu_1 + & \frac{1}{2}\mu_2 \end{cases}$$

§ 3.3

Problem 2, *p.* 202

Y is the number of heads obtained in tossing a fair coin three times, so

$$\mathsf{P}[Y = k] = \binom{3}{k}\left(\frac{1}{2}\right)^3 \qquad k = 0, 1, 2, 3.$$

Then,

$$\begin{aligned} \mathsf{E}[Y^2] &= \sum_{k=0}^{3} k^2 \mathsf{P}[Y = k] \\ &= 0 \times \frac{1}{8} + 1 \times \frac{3}{8} + 4 \times \frac{3}{8} + 9 \times \frac{1}{8} \\ &= 3 \\ \mathsf{E}[Y^4] &= \sum_{k=0}^{3} k^4 \mathsf{P}[Y = k] \\ &= 0 \times \frac{1}{8} + 1 \times \frac{3}{8} + 16 \times \frac{3}{8} + 81 \times \frac{1}{8} \\ &= 16.5 \end{aligned}$$

So

$$\mathsf{Var}[Y^2] = \mathsf{E}[Y^4] - (\mathsf{E}[Y^2])^2 = 16.5 - 3^2 = 7.5$$

Problem 12, *p.* 203

a) Since $\mu = 10$ and $\sigma = 5$, apply Chebychev's inequality,

$$\mathsf{P}[X \geq 20] = \mathsf{P}[X - \mu \geq 2\sigma] \leq \mathsf{P}[|X - \mu| \geq 2\sigma] \leq 1/2^2 = 0.25$$

b) X cannot be binomial. Otherwise, suppose $X \sim \mathsf{Bi}(n, p)$, then $\mu = np = 10$, and $\sigma = \sqrt{npq} = 5$, hence we have $q = 2.5 > 1$, which is impossible.

Problem 14, *p.* 203

Suppose X is the average family income, it is a non-negative random variable.

a) Apply *Markov's Inequality*,

$$\mathsf{P}[X > 50\,000] \leq \frac{\mathsf{E}[X]}{50\,000} = \frac{10\,000}{50\,000} = 0.2$$

b) Apply *Chebyshev's Inequality*,

$$\mathsf{P}[X > 50\,000] = \mathsf{P}[X - 10\,000 > 5 \times 8\,000] \leq \mathsf{P}[|X - 10\,000| > 5 \times 8\,000] \leq \frac{1}{5^2} = 0.04$$

Problem 17, *p.* 204

Calculate the mean and standard deviation for X,

$$\mu = \mathsf{E}[X] = -1 \times \frac{1}{4} + 0 \times \frac{1}{4} + 1 \times \frac{1}{2} = -\frac{1}{4} = 0.25$$

$$\mathsf{Var}[X] = \mathsf{E}[X^2] - (\mathsf{E}X)^2 = 1 \times \frac{1}{4} + 0 \times \frac{1}{4} + 1 \times \frac{1}{2} - \left(-\frac{1}{4}\right)^2 = \frac{11}{16} = 0.6875$$

$$\sigma = \sqrt{\mathsf{Var}[X]} = \frac{\sqrt{11}}{4} = 0.8292$$

Approximate the sum S_{25} as a normal distribution, with mean $\mathsf{E}[S_{25}] = 25\mu = 6.25$, and standard deviation $\mathsf{SD}(S_{25}) = 5\sigma = 4.1458$

a)

$$\mathsf{P}[S < 0] \approx \Phi\left(\frac{-0.5 - \mathsf{E}[S_{25}]}{\mathsf{SD}(S_{25})}\right) = \Phi(-1.6282) = 0.0517$$

b)

$$\mathsf{P}[S = 0] \approx \Phi\left(\frac{0.5 - \mathsf{E}[S_{25}]}{\mathsf{SD}(S_{25})}\right) - \Phi\left(\frac{-0.5 - \mathsf{E}[S_{25}]}{\mathsf{SD}(S_{25})}\right) = \Phi(-1.3869) - \Phi(1.6282) = 0.0310$$

c)

$$\mathsf{P}[S > 0] \approx 1 - \Phi\left(\frac{0.5 - \mathsf{E}[S_{25}]}{\mathsf{SD}(S_{25})}\right) = 1 - \Phi(-1.3869) = 0.9173$$

Problem 18, *p.* 204

Suppose X is winning money, then $p = \mathsf{P}[X = 6] = \frac{5}{38}$, $q = \mathsf{P}[X = -1] = \frac{33}{38}$. Calculate the mean and variance

$$\begin{aligned}
\mu = \mathsf{E}[X] &= 6p - q = -\frac{3}{38} = -0.0789 \\
\mathsf{E}[X^2] &= 36p + q = 5.6053 \\
\mathsf{Var}[X] &= \mathsf{E}[X^2] - \mathsf{E}[X]^2 = 5.5999 \\
\sigma = \mathsf{SD}(X) &= \sqrt{\mathsf{Var}[X]} = 2.3662
\end{aligned}$$

Let $S = X_1 + X_2 + \cdots + X_{300}$, then S has an approximately Normal distribution by *Central Limit Theorem*

$$\mathsf{P}[S > 0] = \mathsf{P}[\frac{S - 300\mu}{\sqrt{300}\sigma} > \frac{0 - 300\mu}{\sqrt{300}\sigma}] = 1 - \Phi(\frac{23.6842}{40.9843}) = 0.2816$$

Problem 24, *p.* 205

Denote by X_k the number obtained from the k^{th} draw, then

$$\mathsf{P}[X_k = 0] = 1/4, \quad \mathsf{P}[X_k = 1] = 1/2, \quad \mathsf{P}[X_k = 2] = 1/4, \quad S_n = \sum_{k=1}^{n} X_k$$

a) Since all the $\{X_i\}$ are independent and indentically distributed,

$$\begin{aligned} \mathsf{P}[S_2 = k] = \mathsf{P}[X_1 + X_2 = k] &= \sum_{i=0}^{2} \mathsf{P}[X_1 = i, X_2 = k - i] \\ &= \sum_{i=0}^{2} \mathsf{P}[X_1 = i]\mathsf{P}[X_2 = k - i] \end{aligned}$$

Then it is easy to get the exact distribution of S_2,

S_2	0	1	2	3	4
$\mathsf{P}[S_2]$	1/16	1/4	3/8	1/4	1/16

b) Calculate $\mu = \mathsf{E}[X_k] = 1$, and $\sigma = \mathsf{SD}(X) = \sqrt{0.5} = 0.7071$.

$$\mathsf{P}[S_{50} = 50] \approx \Phi(\frac{50 + 0.5 - 50}{\sqrt{50 \times 0.5}}) - \Phi(\frac{50 - 0.5 - 50}{\sqrt{50 \times 0.5}}) = \Phi(0.100) - \Phi(-0.100) = 0.0796$$

c) We shall prove that $\mathsf{P}[S_n = k] = \binom{2n}{k}/2^{2n}$ by mathematical induction. For $n = 0$ this statement asserts that $\mathsf{P}[S_0 = 0] = 1$, trivial but true. Suppose the statement holds for all $m \leq n$ for some $n \geq 0$; then

$$\begin{aligned} \mathsf{P}[S_{n+1} = k] &= \mathsf{P}[S_n = k, X_{n+1} = 0] + \mathsf{P}[S_n = k-1, X_{n+1} = 1] + \mathsf{P}[S_n = k-2, X_{n+1} = 2] \\ &= \mathsf{P}[S_n = k]\mathsf{P}[X_{n+1} = 0] + \mathsf{P}[S_n = k-1]\mathsf{P}[X_{n+1} = 1] + \mathsf{P}[S_n = k-2]\mathsf{P}[X_{n+1} = 2] \\ &= \frac{1}{2^{2n}}\left[\binom{2n}{k}\frac{1}{4} + \binom{2n}{k-1}\frac{1}{2} + \binom{2n}{k-2}\frac{1}{4}\right] \\ &= \frac{1}{2^{2n+2}}\binom{2n+2}{k}\left[\frac{(2n+2-k)(2n+1-k)}{(2n+2)(2n+1)} + \frac{k(2n+2-k) \times 2}{(2n+2)(2n+1)} + \frac{k(k-1)}{(2n+2)(2n+1)}\right] \\ &= \frac{1}{2^{2n+2}}\binom{2n}{k} \end{aligned}$$

By Mathematical Induction, the statement is true for every natural number $n \geq 0$. Alternately, we can recognize the distribution of $X_i \sim \mathsf{Bi}(2, \frac{1}{2})$ and hence $S_n \sim \mathsf{Bi}(2n, \frac{1}{2})$, proving the formula.

§ 3.4

Problem 5, *p.* 218

Denote $q_i \equiv (1 - p_i)$ for $i = 1, 2, 3$.

a) The event *It takes Mary more than n tosses to get a head* occurs if and only if Mary's first n tosses are all tails, so the probability is ${q_2}^n$.

b) Each of them must toss only tails in the first n tosses; by independence, the probability is ${q_1}^n {q_2}^n {q_3}^n = (q_1\, q_2\, q_3)^n$.

d) The event *The first person get a head has to toss exactly n times* occurs if and only if all three toss only tails for the first $n - 1$ tosses, but at least one of them tosses a head on the n^{th} toss; its probability is

$$(q_1\, q_2\, q_3)^{n-1}(1 - q_1\, q_2\, q_3) = (q_1 q_2 q_3)^{n-1} - (q_1 q_2 q_3)^n$$

e) One way to compute the probability of the event *the first head is tossed by Mary* is to sum up over the possible numbers of tosses n needed for the first toss to occur; we get

$$\begin{aligned} \mathsf{P}[\text{Mary tosses the first head}] &= \sum_{n=1}^{\infty} \mathsf{P}[\text{No head before } n^{\mathsf{th}} \text{ toss}]\mathsf{P}[\text{Mary head on } n^{\mathsf{th}} \text{ toss}] \\ &= \sum_{n=1}^{\infty} (q_1 q_2 q_3)^{n-1} p_2 \\ &= \frac{p_2}{1 - q_1\, q_2\, q_3} \end{aligned}$$

Problem 10, *p.* 219

a) $\mathsf{P}[X < 2] = 0$, so when $k \geq 2$, if $X = k$, there are only two possibilities, $k - 1$ success followed by a failure, or $k - 1$ failures followed by a success. Let $q = 1 - p$, we have

$$\mathsf{P}[X = k] = p^{k-1}q + q^{k-1}p, \qquad k \geq 2.$$

b)

$$\begin{aligned} \mathsf{E}[X] &= \sum_{k=2}^{\infty} k\mathsf{P}[X = k] \\ &= \sum_{k=2}^{\infty} k[pq^{k-1} + qp^{k-1}] \\ &= (\frac{1}{p} - p) + (\frac{1}{q} - q) \\ &= \frac{1}{pq} - 1 \end{aligned}$$

c)

$$
\begin{aligned}
\mathsf{E}[X^2] &= \sum_{k=2}^{\infty} k^2 \mathsf{P}[X=k] \\
&= \sum_{k=2}^{\infty} k^2 [pq^{k-1} + qp^{k-1}] \\
&= (\frac{1+q}{p^2} - p) + (\frac{1+p}{q^2} - q) \\
&= \frac{p^2+q^2+p^3+q^3}{p^2q^2} - 1 \\
&= \frac{2}{p^2q^2} - \frac{5}{pq} - 1 \\
\mathsf{Var}[X] &= \mathsf{E}[X^2] - \mathsf{E}[X]^2 \\
&= (\frac{2}{p^2q^2} - \frac{5}{pq} - 1) - (\frac{1}{pq} - 1)^2 \\
&= \frac{1}{p^2q^2} - \frac{3}{pq} - 2
\end{aligned}
$$